Research evaluation: bibliometrics and the librarian

Jenny Delasalle
Academic Support Manager (Research)
University of Warwick Library
Tel: +44 (0)24 7615 1275
Email: j.delasalle@warwick.ac.uk

Setting the scene

As Academic Support Manager (Research) at the University of Warwick Library, it is my remit to help shape and move forward the library's support for our researchers. I was formerly the library's E-Repositories Manager and prior to that I was a Service Innovation Officer.[1] I also have experience of working at various other libraries in subject support-type roles and on JISC-funded projects.

My current role was new to the library and new to me in March 2010. I began by investigating researchers' needs: I read published reports, I talked to our Academic Support Librarians, all of whom have subject expertise and knowledge of our departments, and I investigated support for researchers at other university libraries. I also made links with others who support researchers across the university, and consulted researchers themselves about their information skills needs.

Research evaluation arose as a theme relevant to all researchers across the disciplines and it is one in which my previously gained expertise on scholarly communications, citations and publication practices is relevant. Librarians have skills and experience in evaluating the quality of information resources and these are very relevant to researchers whose work is to be evaluated, as well as to those proposing to evaluate research.

In this article I describe some of the work I have been doing on the theme of bibliometrics in the last year. Research evaluation itself includes more than bibliometrics, and researchers seem increasingly engaged by the theme of impact.

However, research evaluation activities by the Higher Education Funding Council for England (HEFCE), which is responsible for the Research Excellence Framework (REF),[2] by university-ranking publishers, by research funders, by heads of department and by institutions as a whole, continue either to incorporate bibliometrics or to consider the role and relevance of bibliometrics.

Reaching researchers

I'd better define 'bibliometrics' before we get any further! A bibliometric is a measure relating to publications. Such a measure could be simply the number of items published. Often, though, we use the term 'bibliometrics' to refer to the measurement and analysis of citations for journal articles published. There are many different measures, and this article discusses some of the most commonly used ones.
I do talk to researchers about 'bibliometrics' itself from time to time, because it is a word that is mentioned in association with the REF and that, plus the mystery of what it means, is a potent attraction, for some! However, many researchers neither know nor care about 'bibliometrics', per se. What they are interested in is what the measures mean to them:

- What do we do about the REF?
- What is an h-index and how do I find mine?
- Getting evidence for a job application/grant application/salary review, etc.
- Considering where to publish: how do I find journal impact factors?
- Benchmarking a department and how it should be done.
- Measurements used in university rankings

When proposing an information skills workshop to a department on bibliometrics, it can be useful that research performance is a controversial topic. It is therefore relatively easy to engage your audience in debate and to display your knowledge and skills and to use the expertise of those taking part in order to get discussions going.

Engaging researchers is crucial to being able to support them. Librarians have expertise which researchers find helpful, but they are often surprised that we can help them; we need to be seen to be relevant to them. At the University of Warwick Library we have a facility called the Wolfson Research Exchange.[3] My colleague, Fiona Colligan, Academic Support Librarian (Research) has been instrumental in developing our model of peer-to-peer support through this facility and its web presence. We're building community resources so that researchers can help each other and can promote the library's offer to each other.

Other than community engagement and contact with separate departments, we offer information skills training. This is delivered through existing programmes that are centrally provided, for staff and for PhD students, as well as web-based guides on both the library site and the Research Exchange site. The unique thing about the Research Exchange website is that the materials are all written by researchers themselves – even though in some cases we provided authors with a brief to write to.

What do researchers need to know?

I find that researchers are very sceptical and like to come to their own conclusions, so I do not try to persuade them of the value of bibliometrics, but I do try to ensure that they are properly informed about what is happening and the potential measures available. It's a topic which seems to be developing and moving on at a relatively fast pace, so it's one in which researchers seem to value support.
From my experience, researchers should know:

- Citation data sources available to them, and about their strengths and weaknesses as sources. Sources include:
  - Web of Knowledge (WoK), including Journal Citation Reports (JCR)[4]
  - Scopus[5]
  - SCImago's Journal and Country Rank[6]
  - Google Scholar[7] and tools to analyse its data, such as Publish or Perish[8] or the Mozilla Firefox extension[9]
  - Other subject specialist sites which include citation data, such as RePEc[10]
- Types of impact factor/citation-based measures for journals available through those sources. Some are based on two, three years or five years of data. Some measures include weightings for citations from highly cited articles. Researchers should know which measure is most appropriate for their needs.
- There is more than one kind of 'h-index'. There is one formula for calculating the actual h-index, but depending on the data set used, a researcher might get a very different score for his/her h-index.
- There are formulae for other kinds of index, such as the g-index and the contemporary h-index. Researchers should consider which index is a measure of the particular aspects

of research that matter most to them or to those measuring their work.
- Other journal rankings are available than purely bibliometrics-based ones: Anne-Wil Harzing's website[11] has some very useful lists of journal rankings in its Journal Quality List, although there is a subject focus on the site towards Economics, Finance, Accounting, Management, and Marketing.
- Which factors are most important to them and their co-authors when writing for publication?

What is the best source of data?

I don't have the answer to this one as it will depend on the discipline of the researcher and what their actual needs are. It might also depend on which source your institution has a subscription to. We have a guide to journal-impact factors on WoK's JCR and the SCImago Journal Rank (SJR) on our Research Exchange website, which I would refer researchers to since it discusses the coverage of Scopus and WoK.[12]

I mostly use Web of Knowledge myself and I find it useful to check Google Scholar's data. WoK's citation data involves more human intervention on its creation than Google Scholar uses, even though Google Scholar apparently indexes more publications. Also, we know that WoK has its historical bias towards English language publications and doesn't cover Arts and Humanities subjects so well, but we don't truly know what Google Scholar indexes from one day to the next!

Google Scholar does index citations from books and other kinds of publication that WoK doesn't currently cover, but Web of Knowledge's data are more likely to be used in the forthcoming REF and they are used in some prestigious university rankings, including that of Times Higher Education.[13]

I would not recommend using data from more than one source in the calculation of an h-index though. It would take a lot of work to calculate manually and would not be easily reproducible by others, and therefore not verifiable.

What is the best type of journal impact factor?

This isn't easy to answer either and it really depends on which factors the researcher is most concerned about. The WoK JCR five-year impact factor[14] is good for comparing journals across disciplines since it takes into account that in some disciplines, a journal's peak rate of citation might occur later than two years after publication, but the SCImago Journal Rank uses a three-year impact factor which will account for more journals' peak citation rates than just two years but not as many as five.

The WoK Immediacy Index is a measure of the number of times a journal article published in a specific year is cited over the course of that same year and it could be useful for researchers who want to publish in a journal from which they may be quickly cited. However, citation practices vary between disciplines, so this kind of measure is very much affected by the discipline of the journal. We have a guide to WoK measures of journals on our Research Exchange website; it includes more information about these.[15]

What other features of a journal should a researcher use to appraise it?

New authors in particular are interested to learn about other features they can use in assessing the quality or suitability of a journal they wish to approach with their article. Here are a few questions that researchers can ask themselves:

- Turnaround time: either for them to reject you so that they can approach another journal, or from acceptance to publication. It might be of paramount importance that your work is published quickly. You can sometimes find this kind of information on a journal's home page or you might need to contact them to ask about such lead times.
- Editors and peer reviewers: are these people advertised on the journal's website and are they prestigious in your field?
- Audience: who reads this journal, and who are you and your co-authors trying to reach? One clue to this might be which websites index the content of your journal: are these sites the main places where researchers in your discipline will be searching?
- Copyright/author agreement offered: if the journal allows you to deposit an Open Access (OA) version straight away then it might not be so crucial how quickly the journal itself publishes your work. It might be important to you that your article is made available on open access: the Directory of Open Access Journals (DOAJ)[16] lists open access journals, and SherpaRomeo[17] is particularly useful for finding links to journals' open access policies.

- Are there other features of the journal that you like, such as podcasts,[18] a readers' comments space, a view of statistics on number of accesses to your article and such like? [19]

What is the h-index?

This is a measure invented by a physicist, Jorge E. Hirsch, and is an algorithm intended to indicate the quality and sustainability of research output. It is calculated using the number of publications and the number of citations per publication. A researcher or group with an index of *h* has published *h* papers each of which has been cited by others at least *h* times.

An h-index of 4 means there are 4 published papers each with at least 4 citations.

This can be shown graphically, as a list:

Publication 1	20 cites
Publication 2	18 cites
Publication 3	11 cites
Publication 4	7 cites
----------	h-index: 4
Publication 5	4 cites
Publication 6	3 cites
Publications 7, 8, 9, 10	0 cites

If our example researcher were to reach an h-index of 5 then s/he would need publication no 5 to gain one more citation.

What about these other measures that are like the h-index?

I like two other measures of researchers' work:

- Egghe's g-index which, aims to improve on the h-index by giving more weight to highly-cited articles.[20]
- The 'Contemporary h-index', which aims to improve on the h-index by giving more weight to recent articles and therefore rewards academics who maintain a steady level of activity[21]

But again, it all comes back to what you wish to measure as to which kind of score is going to be best suited to your needs.

Boosting citations

A question that engages many researchers is: 'Are there tactics which boost the number of citations to my articles?' To which my answer is: 'Possibly!' Here are a few of the possibilities that I share with researchers:

- OA publishing supporters claim that OA articles are cited more highly.[22]
- There are many routes to OA, including depositing your article in our institutional repository.[23]
- The more ways there are of someone discovering and reading your work, the more likely it is that they will also cite it.
- Review papers accrue more citations than those discussing original research: you will notice that journals specialising in these papers have the highest journal impact factors.
- Consider writing fewer, larger papers so that there is more content in a single article to be cited.
- Self-citations aren't always stripped out of analyses, and in any case, they can lead to others discovering and therefore citing your work.
- Monitor citations of your work: set up alerts on citation tracking websites such as WoK, so that you get an email every time someone cites your article.

In conclusion

It's not only researchers who need to know all of this of course. Often, university administrators also need a good grasp of bibliometrics and of what it is they are trying to measure and why. Universities wishing to score well in the REF or in rankings might want to encourage their researchers to boost their h-index scores: both measures are based on citations, so in a way it makes sense. However, it's not that simple because the way they use the data might be very different.

Research performance, particularly when it involves measures of publications, is a very complicated arena and I still do not think I have got to the bottom of bibliometrics. Having said that, I think I've got a good enough grasp to inform a lot of researchers in a way that helps them to make sure that they are able to present their own research in the best light and to question the methodology of any research evaluation exercises which use bibliometrics, in an informed way.

Further reading and resources

- I recommend the bibliometrics toolkit at http://www.ndlr.ie/myri/. It's a great tuto-

rial, even if you don't download and customise it.

- Read all about the REF bibliometrics pilot at http://www.hefce.ac.uk/research/ref/biblio/
- The University of Southampton's Library website has excellent materials on bibliometrics:http://www.soton.ac.uk/library/research/bibliometrics/index.html
- A lot more detail about what Thomson Reuters are doing can be found on their website, and I find their training section, especially the recorded training particularly helpful: <http://thomsonreuters.com/products_services/science/training/jcr/
- Likewise, the Elsevier website has a useful section on bibliometrics: http://www.elsevier.com/wps/find/editorshome.editors/biblio

Notes

1 J. Delasalle, 'A week in the life', SCONUL *Focus*, 40 (Spring), 2007, pp. 12-14; online at http://www.sconul.ac.uk/publications/newsletter/40/4.pdf
2 See http://www.hefce.ac.uk/research/ref/
3 See http://go.warwick.ac.uk/researchexchange
4 http://wok.mimas.ac.uk/
5 http://www.scopus.com/
6 http://www.scimagojr.com/
7 http://scholar.google.co.uk/
8 http://www.harzing.com/pop.htm
9 https://addons.mozilla.org/en-US/firefox/addon/scholar-h-index-calculator/
10 http://repec.org/
11 http://www.harzing.com/
12 See http://go.warwick.ac.uk/researchexchange/topics/gd0055
13 See http://www.timeshighereducation.co.uk/world-university-rankings/
14 See http://thomsonreuters.com/products_services/science/free/essays/impact_factor/
15 See <http://go.warwick.ac.uk/researchexchange/topics/gd0054
16 http://www.doaj.org/
17 http://www.sherpa.ac.uk/romeo/
18 The Sage journal, Human Relations features podcasts: http://hum.sagepub.com/
19 Many of the PLoS journals offer a readers' response mechanism and offer publicly visible metrics: http://www.plos.org/journals/index.php
20 See http://en.wikipedia.org/wiki/G-index
21 Anne-Wil Harzing discusses these kinds of measure at http://www.harzing.com/pop_hindex.htm
22 On this topic, I refer researchers to the Open Access Bibliography: http://www.digitalscholarship.org/oab/oab.htm
23 At Warwick our repository is WRAP: http://wrap.warwick.ac.uk